

Structure and lipid transport mechanism of a StAR-related domain

Yosuke Tsujishita and James H. Hurley

The steroidogenic acute regulatory protein (StAR) regulates acute steroidogenesis in the adrenal cortex and gonads by promoting the translocation of cholesterol to the mitochondrial inner membrane where the first step in steroid biosynthesis is catalyzed. StAR-related lipid transfer (START) domains occur in proteins involved in lipid transport and metabolism, signal transduction, and transcriptional regulation. The 2.2 Å resolution crystal structure of the START domain of human MLN64 reported here reveals an α/β fold built around a U-shaped incomplete β -barrel. The interior of the protein encompasses a $26 \times 12 \times 11$ Å hydrophobic tunnel that is large enough to bind a single cholesterol molecule. The StAR and MLN64 START domains bind 1 mole of ^{14}C cholesterol per mole of protein *in vitro*. Based on the START domain structure and cholesterol binding stoichiometry, it is proposed that StAR acts by shuttling cholesterol molecules one at a time through the intermembrane space of the mitochondrion.

The steroidogenic acute regulatory protein (StAR) related lipid transfer (START) domains are 200–210 amino acid motifs that occur in a remarkably wide range of proteins involved in diverse cell functions^{1,2} (Fig. 1). StAR mobilizes cholesterol to the inner mitochondrial membrane of steroidogenic cells in the adrenal cortex and gonads^{3–11}. This cholesterol transfer process is the principal control point for regulation of steroidogenesis by adrenocorticotropin (ACTH) and other hormones acting through the adenyl cyclase and Ca^{2+} pathways. The closest homolog to StAR is MLN64, with 35% sequence identity between the two in their START domains. MLN64 may play a StAR-like role in promoting steroidogenesis in placenta and brain, and is overexpressed in some breast carcinomas^{12,13}. MLN64 is also expressed in some nonsteroidogenic tissues, where cholesterol is oxidized to hydroxysterols by mitochondrial P450 enzymes in the first step of bile acid synthesis.

Cholesterol transfer is only one of the biological processes in which START domain proteins are implicated. Among other examples of START domain proteins, phosphatidylcholine (PC) is specifically mobilized by PC transfer protein (PC-TP). A human putative acyl-CoA thioesterase contains a START domain, suggesting that it has a role in lipid metabolism¹. The signal transducing proteins p122-RhoGAP^{1,14}, Goodpasture antigen binding protein^{2,15}, and several plant and *Caenorhabditis elegans* pleckstrin homology (PH) domain proteins² all contain START domains. Finally, plant transcription factors of the homeodomain and leucine zipper containing (HD-Zip) and Gl2 families^{1,16} contain START domains, suggesting they may be nuclear receptors for an as yet unidentified class of plant-specific lipophilic messengers¹.

The structure determination of START domains was initiated for two main reasons. By systematically solving the structures of intracellular signaling domains, including

the START domain, it is hoped that new insights will be obtained into the wide range of signaling pathways in which these domains participate. For the START domain in particular, there has been an intense interest over the past five years in the mechanism by which StAR regulates acute steroidogenesis^{7,8,10,11} — the rapid production of steroids by the adrenal cortex and gonads in response to ACTH stimulation.

Steroid biosynthesis begins with the cleavage of the cholesterol side chain by the cytochrome P450 side chain cleavage enzyme (P450_{sc}) to yield pregnenolone. P450_{sc} resides on the matrix side of the inner mitochondrial membrane. The rate limiting step in pregnenolone synthesis is the delivery of cholesterol from the outer mitochondrial membrane to the inner mitochondrial membrane by StAR. The role of StAR in this process is unique, and other types of cholesterol transporters, such as sterol carrier protein-2 (SCP2), cannot replace it.

Cholesterol transport by StAR has a number of unusual features. The function of StAR in steroidogenesis depends on new protein synthesis, yet the degradation of the StAR protein is much slower than its functional inactivation^{8,10–11}. Although StAR promotes the transport of cholesterol between the outer and inner mitochondrial membranes, transfected StAR mutants that cannot be imported into the mitochondrion still support steroidogenesis in cells¹⁷. Recombinant protein lacking the mitochondrial import sequence also promotes steroid production in

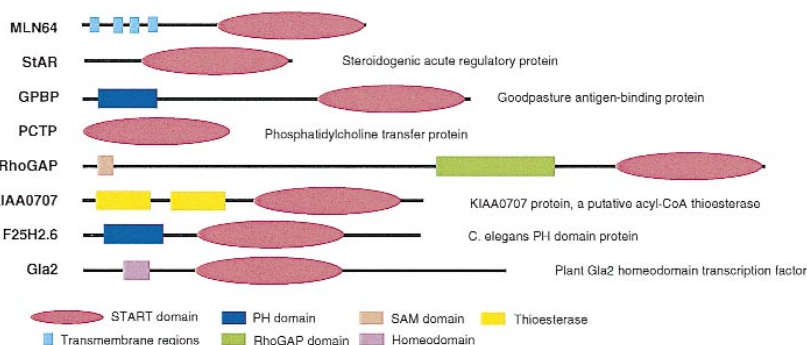


Fig. 1 Representative START domain containing proteins, based on information obtained from the SMART database². RhoGAP, Rho GTPase activating protein.

Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0580, USA.

Correspondence should be addressed to J.H.H. email: jh8e@nih.gov

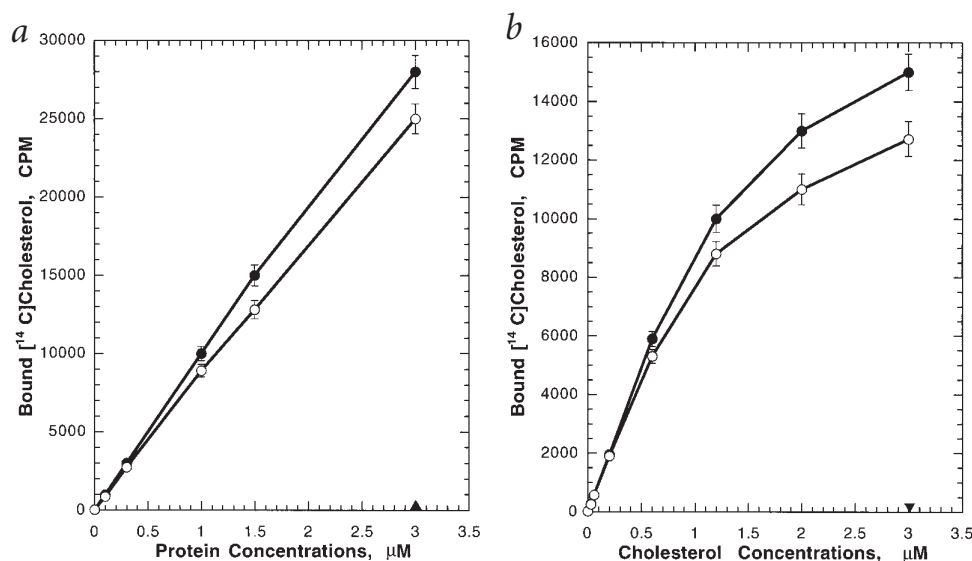


Fig. 2 Cholesterol binding to MLN64 and StAR. **a**, Indicated concentrations of His₆-tagged MLN64 (empty circles), StAR (filled circles), and inactivated MLN64 (triangle) were incubated with 3 μM [¹⁴C] cholesterol for 1 h at 37 °C. Radioactivity of [¹⁴C] cholesterol bound to eluted proteins was quantified. **b**, Indicated concentrations of [¹⁴C] cholesterol were incubated with 1.5 μM His₆-tagged MLN64 (empty circles) or StAR (filled circles) for 1 h at 37 °C, and [¹⁴C] cholesteryl oleate was incubated with 1.5 μM His₆-tagged MLN64 (inverted triangle).

0.9:1 for StAR-START and 0.8:1 for MLN64-START.

The saturable binding at approximately unit stoichiometry suggests that bind-

ing occurs at a specific site, and is not the result of nonspecific aggregation of protein and cholesterol. Further, cholesteryl oleate does not bind to either START domain, showing that binding is specific and governed by steric constraints. Finally, MLN64-START that has been inactivated by seven days storage at 25 °C does not bind cholesterol. Taken together, these observations show that cholesterol binds to both the START domains of StAR and MLN64 with similar affinity and stoichiometry, and that binding occurs at a single specific site.

Overall structure of the MLN64 START domain

The structure of MLN64-START was determined to 2.2 Å resolution by MAD phasing using a selenomethionyl derivative of the mutant F388M (Table 1, Fig. 3a,b). The structure is of the α/β type, consisting of a nine-stranded twisted antiparallel β-sheet and four α-helices (Fig. 4a,b). The β-sheet has strand order 123987654 and forms a U-shaped unclosed β-barrel. Two Ω-loops are inserted between β5-β6 and β7-β8. In the most unusual feature of the β-sheet, strand β7 stops short of its neighbors β6 and β8 to begin the second Ω-loop. Gly 381, Gly 384, and Pro 385 facilitate the formation of two kinks in β8. Preceding the kinks, β8 forms a short section of parallel β-sheet with β6.

The most striking feature of the START domain structure is a predominantly hydrophobic tunnel extending 26 Å, nearly the entire length of the protein (Fig. 5a,b). There are two openings of ~3 Å in diameter at either end, but the central chamber of the tunnel has a cross section as wide as 11 Å × 12 Å and a volume of ~1,900 Å³. The floors and walls of the tunnel are formed by the interior face of the incomplete β-barrel, while the roof is formed by helices α3 and α4 and by the Ω-loop inserted between β5-β6. The second Ω-loop, inserted between β7-β8, buttresses the first. The kinks in β8 result in a short branched channel that forks away from the central chamber of the tunnel. The tunnel is not completely hydrophobic, and contains a buried charged pair between Asp 332 and Arg 351. The refined structure includes 11 bound solvent molecules within the tunnel. Given the resolution of this structure, this probably represents only a part of the ordered solvent in the tunnel. The electron density, however, does not contain any connected features suggestive of a bound lipid or other large molecules.

The START domain structure revealed an unexpected similarity to that of the birch pollen allergen Bet v 1 (ref. 20), based on searches of the nonredundant protein structure database²¹. The

a cell-free system containing mitochondria¹⁸. This mutant is essentially identical to the isolated START domain of StAR.

Taken together, these observations have led to a model in which StAR acts on the outer face of the outer mitochondrial membrane, rather than within the intermembrane space^{17,18}. In this model, the import of StAR into the mitochondrion inactivates it. The mechanism by which StAR can carry out cholesterol transport from the cytoplasmic face of the outer membrane remains unclear. The search for this mechanism has guided much of the experimentation in the StAR field over the past three years. One model invokes cholesterol desorption from the outer face of the outer membrane¹⁸, while another suggest that StAR is active as a molten globule associated with the mitochondrial protein import channel¹⁹. As yet, neither these nor any other mechanisms have been conclusively established as correct.

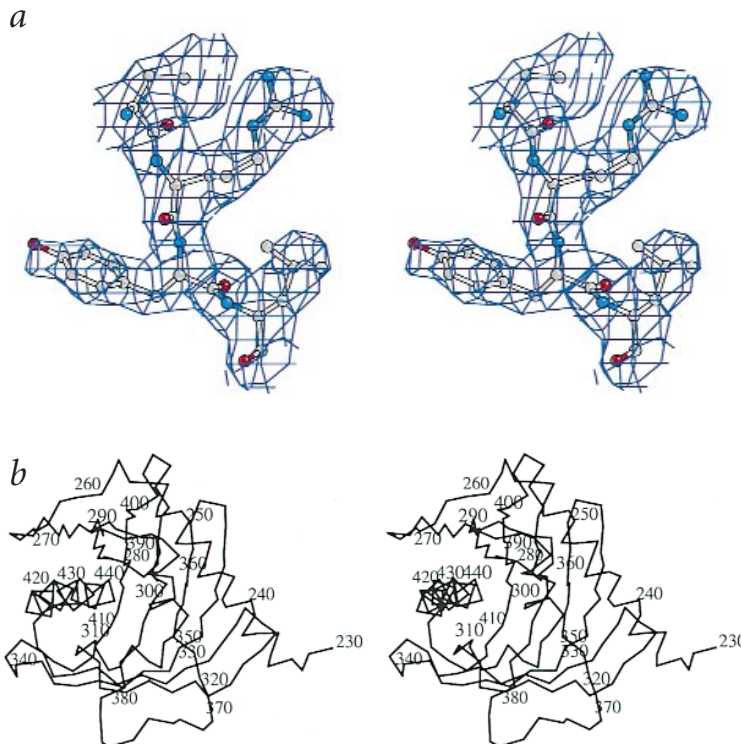
Here, the structure of the START domain of MLN64 is presented. StAR and MLN64 share high (35%) sequence identity within their START domains and have very similar biological functions. The START domain of StAR (StAR-START) was prone to aggregation at high protein concentration and was unsuitable for crystallization. The START domain of MLN64 (MLN64-START), on the other hand, was reasonably soluble and did not aggregate in solution at protein concentrations required for crystallization. MLN64-START is sufficiently similar to StAR-START that we feel it is reasonable to make inferences about one from the structure of the other. To test this assumption, the cholesterol binding properties of the two START domains were characterized and found to be nearly indistinguishable.

Cholesterol binding by START domains

Cholesterol is insoluble in aqueous solution, making equilibrium binding analysis in solution impossible. To date no data have been reported on direct binding of cholesterol to START domain proteins. To determine whether cholesterol binds directly to the START domains of StAR and MLN64, His₆-tagged START domains were titrated with cholesterol dissolved in ethanol (Fig. 2). Cholesterol binds to both domains in a dose-dependent and saturable manner. Because of the insolubility of cholesterol, a true equilibrium between bound and free cholesterol is not established under these conditions, and the equilibrium constant cannot be calculated. At 1.5 μM protein, essentially all of the cholesterol added appears to bind to the protein until saturation is achieved at a stoichiometry of

articles

Fig. 3 Structure determination. **a**, Solvent flattened MAD electron density map at 2.5 Å contoured at 1.0 σ in the vicinity of residues 358–361. The final model is superimposed on the maps. **b**, Stereo C α trace of the MLN64-START domain. Every 10th residue is labeled.



two structures are superimposable over 122 residues with a root mean square (r.m.s.) deviation of 2.8 Å, despite sharing only 9% sequence identity. Weaker similarities to phosphoglucomutase and many other complete and incomplete β -barrel structures were also detected. Bet v 1 contains a large open cleft in place of the tunnel in START domain, and is missing the two Ω -loops that form part of the tunnel roof. The functions of Bet v 1 and its cleft are unknown. Other allergens belonging to the lipocalin family are known to bind lipids in the interior of β -barrel structures²², although their structural similarities to START and Bet v 1 are not strong enough to demonstrate a conclusive homology.

Structural basis for congenital lipid adrenal hyperplasia

In 1996, the international congenital adrenal hyperplasia (CAH) consortium reported an analysis of genetic data from CAH patients in 10 different countries⁹. All of the mutations map to the gene encoding StAR. Missense, frameshift, and stop codon mutations in the START domain of StAR can be mapped onto the MLN64-START structure, and their structural consequences rationalized. These mutations (StAR numbering) include E169G, E169K, R182L, A218V, L275P, and Δ R272.

One important finding in this study is that three of the most common sites of mutations — Glu 169, Arg 182, and Ala 218 — are closely clustered in the tertiary structure (Fig. 4b), even though they are dispersed in the primary sequence (Fig. 6). Glu 169 is the StAR counterpart of MLN64 Asp 332, one of the two charged residues within the tunnel. Replacement of this Glu residue with an uncharged or positively charged residue would in theory lead to a destabilizing charge imbalance in the tunnel, explaining its loss of function phenotype *in vivo*, and its misfolding phenotype *in vitro*²³. Furthermore, the acidic side chain of the Glu or Asp could be involved in specific ligand binding, perhaps with the 3-hydroxyl group of cholesterol. StAR Arg 182 corresponds to Arg 345 in MLN64. This residue makes hydrogen bonding interactions to main chain groups on Ω -loop 2, thereby stabilizing its conformation. The aliphatic part of the Arg side chain forms part of the hydrophobic wall of the tunnel. StAR Ala 218 corresponds to Gly 381 in MLN64, which is at a location where β 8 kinks and switches partners between β 6 and β 7. The kink in β 8 appears to play a key role in stabilizing the branched structure of the tunnel. Gly 381 is tightly packed against tunnel-lining residues from β 6 and β 9. The introduction of a more bulky side chain might therefore perturb the conformations of at least three different secondary structural elements (β 6, β 8 and β 9) involved in forming the walls of the tunnel. Leu 275 and Arg 272 map to the C-terminal helix α 4, which has a major role in the formation of the roof over the tunnel. Presumably their mutation to Pro or deletion would substantially destabilize the helix.

Lipid binding and transport mechanisms

Cholesterol transport by StAR-START has been directly demonstrated *in vitro*²⁴ and *in vivo*^{3–5}, but the mechanism of transport has

been unclear. There is also good reason to believe MLN64-START is a cholesterol transporter, since overexpressed MLN64-START strongly promotes steroidogenesis when transfected into p450scc-expressing COS cells¹³. Indeed, the isolated START domain of MLN64 is more potent in promoting steroidogenesis than intact MLN64 (ref. 13). This is most likely because intact MLN64, which contains four predicted transmembrane helices, is membrane localized and cannot diffuse through aqueous compartments. It has been suggested that intact MLN64 might be activated *in vivo* by the proteolytic cleavage of its START domain from the transmembrane portion of the protein¹³. Also, the results reported here indicate that these domains directly bind a single molecule of cholesterol, suggesting that the transport process involves direct binding of cholesterol monomers.

The START domain of PC-TP directly binds and transports PC²⁵, but has not been reported to bind or transport cholesterol. It is not surprising that similar structural motifs can bind these two lipids. Cholesterol and PC are similar in size, and neither carries a net charge. The nonspecific lipid transport protein SCP2, another hydrophobic tunnel protein²⁶, binds both compounds with similar affinity. These START domain proteins differ markedly from SCP2, however, in that they are selective for one, but not both, lipids.

The tunnel in the MLN64-START structure is just large enough to accommodate one molecule of cholesterol (Fig. 7). The size and hydrophobicity of the tunnel in the MLN64-START structure strongly suggest that START domains bind a single molecule of a large lipophilic compound within the hydrophobic tunnel. In agreement with the structural inference, the best estimate is that both START domains bind a single molecule of cholesterol. This model is supported for PC-TP as well, since chemical modification of PC-TP previously identified tunnel-lining residues corresponding to MLN64-START strand β 9 as part of the PC binding site²⁵. This strand forms part of one wall of the tunnel.

The importance of the tunnel for function is highlighted by the observation that disease-linked mutations map either to the interi-

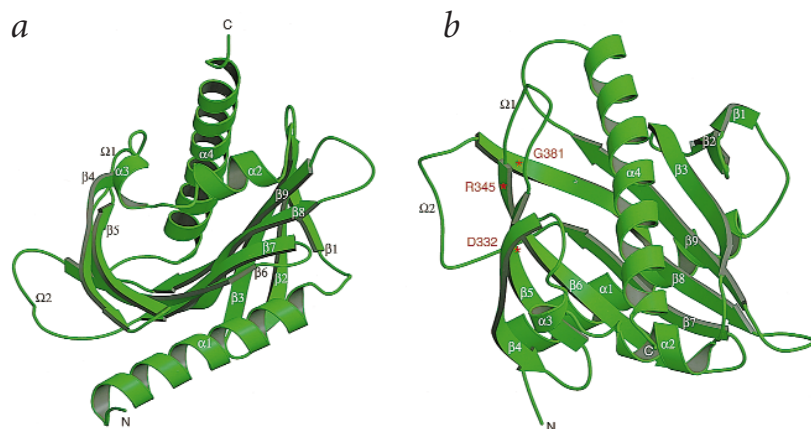


Fig. 4 Overall structure of the START domain of MLN64. **a**, Ribbon diagram of MLN64-START. **b**, A different view of MLN64-START. This view is related to that in (a) by a 90° rotation about the x-axis. The three most common missense mutations responsible for congenital adrenal hyperplasia are marked in red.

or of the tunnel or to $\alpha 4$, which forms part of the tunnel roof. Few residues in these key regions are absolutely conserved (Fig. 6), consistent with both known and postulated differences in the ligand specificity of different START domains. Cholesterol can be modeled in the tunnel with reasonable stereochemistry. Its 3-hydroxyl group might interact with the buried Arg 351 and the rest of the molecule could make extensive hydrophobic contacts with the protein (Fig. 7).

The location of the binding site in a nearly solvent inaccessible tunnel suggests to us a straightforward mechanism for lipid transport by START domain proteins. In this model, cholesterol, PC, or similar large hydrophobic compounds enter the tunnel upon transient opening of the roof by conformational changes in the Ω -loops, pivoting of the C-terminal helix $\alpha 4$, or both. The hydrophobicity of the surface surrounding Ω -loop 1 (Fig. 5b) suggests that the loop could be the membrane docking site for lipid loading and unloading. Once sequestered in the tunnel, these lipids could be transported through the cytoplasm or other aqueous compartments to their destination.

Lipid specificity of START domains

START domains are distinguished from some other groups of lipid transporters by their high degree of specificity. By mapping the aligned START domain sequences onto the crystal structure, we can now begin to identify the key determinants for specificity. The sequence conservation of tunnel wall residues was examined for two of the better understood functional subgroups of START domains, the cholesterol binding domains of StAR and MLN64, and the PC binding domain of PC-TP. Two positions stand out as conserved within, but not between, the subgroups. Met 307 of MLN64 is conserved in StAR, but replaced by an Arg in PC-TP and several other START domains (Fig. 6). Asn 311 of MLN64 is also conserved in StAR, but is replaced by an Asp in PC-TP and either an Asp or Glu in most other START domains. The side chains of these residues are 5 Å apart in the MLN64-START structure, so minor conformational

adjustments in PC-TP-START could place the corresponding Asp and Arg close enough to form a salt bridge. The replacement of the Met/Asn pair of StAR by an Arg/Asp pair increases the polarity of the tunnel in PC-TP. This is consistent with PC-TP binding a zwitterionic lipid rather than an almost completely hydrophobic lipid. Since most START domains also contain the Arg/Asp pair, these may recognize zwitterionic lipids instead of sterols.

Mechanism for cholesterol translocation

The crystal structure and cholesterol binding data obtained here support are both consistent with a model in which StAR binds and solubilizes one cholesterol molecule at a time and shuttles it across the intermembrane space of the mitochondrion. This model is consistent with the report that StAR-START is a sterol transfer protein *in vitro*²⁴. The intracellular transport of numerous lipids, including fatty acids²⁷ and phosphatidylinositol²⁸ and the nonspecific transport of cholesterol and phospholipids by SCP2 (ref. 26), all use a similar mechanism in which lipid monomers are solubilized by the carrier protein. The structures of all of these transporters are known, and they all contain internal hydrophobic cavities, clefts, or tunnels. Compared to these other proteins, StAR appears to be a classic lipid transporter.

The observation that the START domain has a classic lipid transporter-like structure raises new objections to at least one current model of StAR function. Bose *et al.*¹⁹ proposed that StAR functions as a molten globule localized at the outer mitochondrial membrane. The StAR molten globule is observed at pH values below 4.0, however, while the pH values of the cytosol and the intermembrane space are fairly near neutral. In light of the crystal structure, the model that proposes that StAR is active only as a molten globule seems implausible. If StAR has evolved a fold ideally suited to transporting lipids, or binds cholesterol with unit stoichiometry and high affinity in its folded form, it seems unlikely that it would then function *in vivo* only in its molten globule form.

The intermembrane shuttle hypothesis seems the most parsimonious model consistent with the structural and *in vitro* data. However, this model does not explain some of the *in vivo* observa-

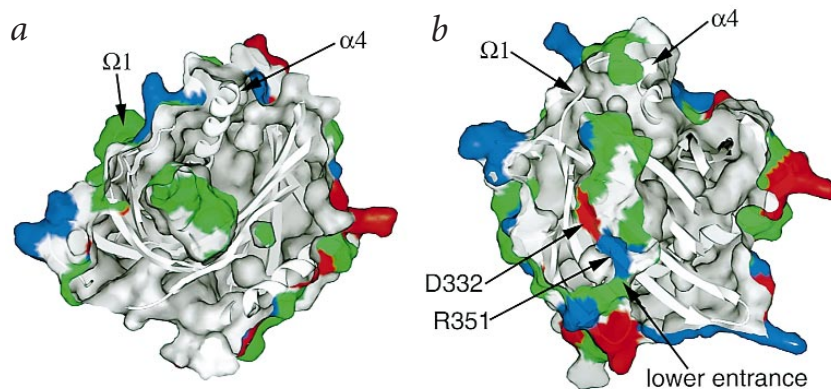


Fig. 5 Cutaway view of the molecular surface of the START domain. **a**, The lipid binding tunnel is formed by the U-shaped β -barrel and the helix $\alpha 4$. This view is the same as that in Fig. 4a. **b**, An alternative view of the lipid binding tunnel revealing the entrance to the tunnel. This view is the same as that in Fig. 4b. The protein molecular surface is colored by residue type: green for hydrophobic, red for acidic, blue for basic and white for uncharged polar. This figure was prepared with Spock⁴⁴.

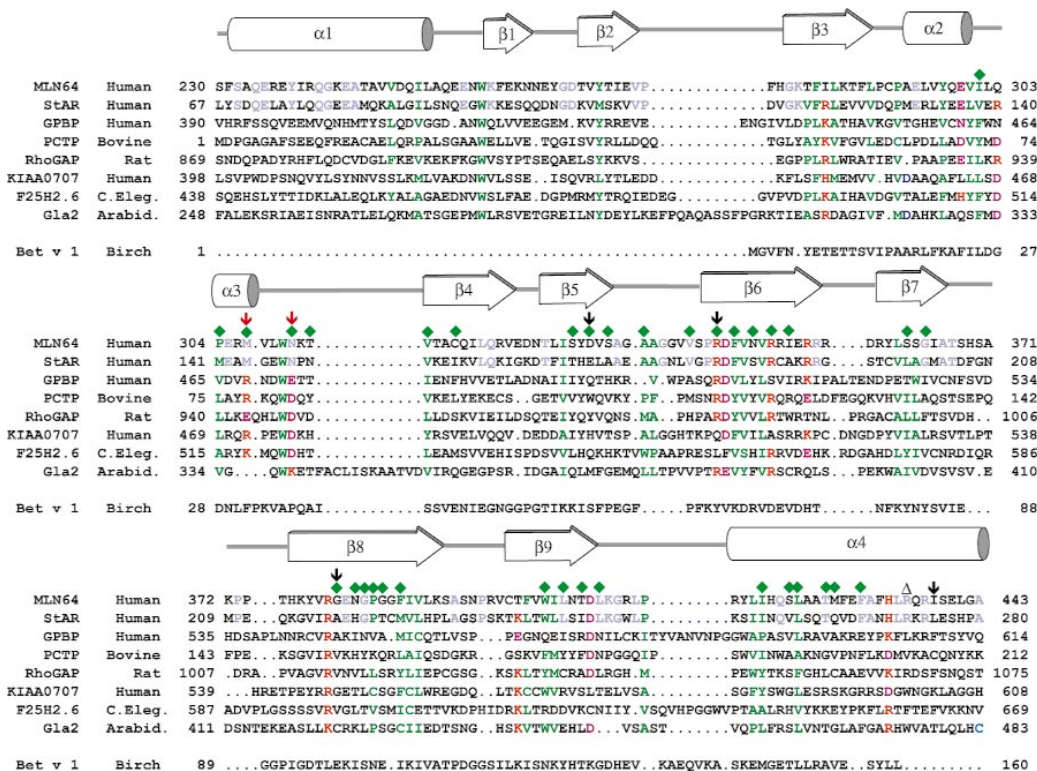


Fig. 6 Structure based alignment of representative START domain sequences. Multiple sequence alignments^{1,2} were modified to eliminate gaps within secondary structural elements where possible. Sequence identities between MLN64 and StAR are highlighted in light blue, conserved charged residues are colored red, and conserved hydrophobic residues are highlighted in green. Residues directly contributing to the tunnel are marked with green diamonds. STAR missense mutations that cause congenital adrenal hyperplasia are marked with black arrows. Putative lipid specificity determinants are marked with red arrows. The single residue deletion in congenital adrenal hyperplasia is marked with a black triangle.

tions. The intermembrane shuttle mechanism is inconsistent with the notion that StAR acts on the outside of the outer mitochondrial membrane^{17,18}. However, the experiments on which this model is based are not immune to alternative interpretations. A key set of experiments involved the transfection of COS cells with N-terminally truncated StAR that was not imported into the mitochondrion at detectable levels. These experiments involved high levels of overexpression, and it is conceivable that a low level of truncated StAR was imported into the mitochondrion through cryptic import sequences or intrinsic leakiness of the import apparatus. In the bacterial protein export system, it has been documented that proteins without signal sequences can be translocated at reduced rates under some conditions²⁹. The same could have occurred in the cell-free experiments reported by the same investigators. It is conceivable that a low level of StAR localized within the mitochondria could be responsible for the observed cholesterol transfer but was not detected in previous studies because of the low relative abundance compared to the very high levels of overexpressed protein in the cytoplasm. Indeed, an electrochemical gradient across the mitochondrial membrane is required for StAR-mediated steroidogenesis⁸, which is consistent with a requirement for protein import.

Another possible objection to the intermembrane shuttle model is raised by the observation that only newly synthesized StAR is functional in cholesterol transport⁸. The shuttle model does not explain the apparent rapid inactivation of mitochondrial StAR. The direct coupling of StAR's import to its inactivation, as postulated in current models, does offer one explanation. In view of the repeated transient interactions with the inner and outer mem-

branes that would be required by the shuttle model, at least one alternative possibility suggests itself. During membrane interaction, a partial opening of the structure is clearly necessary for lipid uptake and offloading because the tunnel openings (Fig. 5b) are too small to admit a cholesterol molecule. During this process, exposed hydrophobic regions of the proteins might form irreversible interactions with the membrane. Once tightly adsorbed to the membrane in this way, StAR would no longer be able to shuttle back and forth. The StAR protein might persist in this state for an extended period before being degraded by proteases. Further experimentation is required to determine whether this is the case. While neither the shuttle nor other models explain all of the data, we believe that the direct shuttle mechanism offers the simplest interpretation of the full body of structural, *in vitro*, and *in vivo* observations than the available alternatives.

Regulation of StAR by phosphorylation

StAR was first identified as a phosphoprotein whose expression was stimulated by ACTH. Both the expression and phosphorylation of StAR are promoted by cAMP. Phosphorylation of Ser 195 increases the activity of StAR *in vivo* as demonstrated by analysis of phosphorylation site mutants³⁰. The MLN64 counterpart of this Ser is Asp 358. Asp 358 is located on the protein surface and does not participate in the hydrophobic tunnel. The location on the surface of the START domain suggests that this residue does not play any part in regulating the overall conformation of the domain nor does it have a direct role in regulating cholesterol binding. Ser 195 phosphorylation fails to activate the isolated StAR-START domain, however. This finding, together with the structural observation,

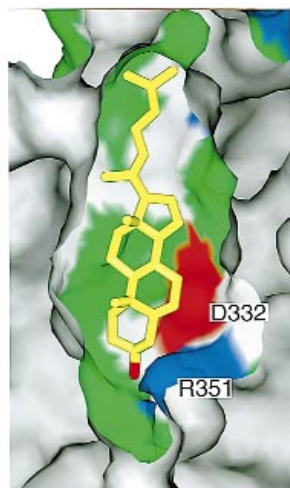


Fig. 7 Model of cholesterol bound in the START domain tunnel. Cholesterol was docked using Spock. Cholesterol is colored yellow (carbon atoms) and red (oxygen atom), and the protein molecular surface is colored as in Fig. 5. The model is intended to show how cholesterol might bind to the StAR-START and MLN64-START domains and to illustrate the dimensions of the tunnel.

implies that phosphorylation of the StAR-START domain regulates the interaction of the N-terminal mitochondrial targeting sequence with the START domain, rather than the ability of StAR to bind cholesterol. This in turn might regulate the rate of import of StAR into mitochondria.

Mechanisms for signal transduction

Many START domains almost certainly have roles in transcriptional regulation and signal transduction, rather than lipid transport¹. It has been suggested that START domains might be membrane targeting motifs^{1,31}. Indeed, this seems to be the most common role for lipid binding motifs in signal transduction. Protein kinase C homology-1 (C1), protein kinase C homology-2 (C2), FYVE, and PH domains are all widely distributed domains whose principal function is to bind membrane lipids and target proteins to specific subcellular loci³¹. The structures of these and other membrane targeting motifs are quite distinctive. Lipid headgroups bind specifically to a pocket or groove on the protein surface, which is surrounded by exposed nonspecific interaction surfaces that may be basic, hydrophobic, or both. With a few exceptions, these structures seem to be designed to recognize specific lipids embedded in phospholipid bilayer aggregates, as opposed to lipid monomers.

In contrast to START, none of these membrane targeting domains contains a buried hydrophobic lipid binding site. The START architecture appears well designed to bind and solubilize lipid monomers, but not to form strong interactions with cell membranes. Structural considerations argue against a membrane-targeting role for START domains. Consistent with these structural inferences, the isolated StAR-START is diffusely distributed throughout the cytosol¹⁷.

The domain structures of signaling and transcriptional regulatory proteins containing the START domain argues against it having a role purely as a lipid transporter. It is difficult to rationalize the presence of homeodomains or RhoGAP catalytic domains, for example, in a protein whose sole function is lipid transport. The nuclear retinoid³², steroid^{33,34}, and peroxisome proliferator-activated³⁵, and related receptors provide a relevant analogy, although their helical ligand binding domains have no homology to the START domain. Transcriptional activation by these nuclear receptors involves ligand-induced conformational changes in the C-terminal helix and an Ω -loop. These structural units

form a hydrophobic lid over their respective ligands that is reminiscent of the roof of the START domain tunnel. It is postulated that the START domains of proteins such as p122-RhoGAP and the Glu2 homeodomain are receptors for specific lipophilic signaling molecules that act by inducing conformational changes in the Ω -loops and the C-terminal helix $\alpha 4$.

The sequences and structural contexts of different START domains show a high degree of divergence. The repertoire of START domain ligands is likely to extend beyond the known binding of sterols and phospholipids. Further investigation to identify the lipophilic regulatory molecules that bind this class of START domains should lead to important new insights into mammalian signal transduction, transcriptional regulation of plant development, and other areas of biology where START domains are represented.

Methods

Protein expression and purification. DNA encoding human MLN64 residues 216–444 was amplified by PCR using an expressed sequence tag (IMAGE 993717) as a template and subcloned into the *NcoI* and *EcoRI* sites of a modified pET22b vector³⁶, which encodes a N-terminal His₆-tag that is cleavable with TEV protease. F388M was introduced by PCR to provide an additional Met for selenomethionyl derivitization. The His₆-tagged protein was expressed in *Escherichia coli* B834 (DE3) cells (Novagen). Selenomethionyl MLN64 was purified from lysate using Ni-NTA resin (Qiagen) and cleaved by incubation with TEV protease (Life Technologies). The cleaved protein was separated from the His₆-tag and His₆-tagged TEV protease by Ni-NTA resin. Protein yield was 20 mg l⁻¹ bacterial culture. Purified protein was dialyzed in 150 mM NaCl, 20 mM Tris-HCl, pH 8.0, and 10 mM DTT and concentrated to 8 mg ml⁻¹.

Cholesterol binding assay. Binding assays were modified following the procedure of Okamura *et al.*³⁷. The purified His₆-tagged START domains of MLN64 and StAR were incubated with [¹⁴C] cholesterol in a buffer containing 150 mM NaCl, 50 mM Tris-HCl, pH 7.4, and 2% (v/v) ethanol for 1 hour at 37 °C. His₆-tagged proteins were immobilized on Ni-NTA resin, washed three times to remove unbound cholesterol and eluted. The radioactivity of [¹⁴C] cholesterol bound to the eluted proteins was quantified with a LS6500 scintillation counter (Beckman).

Crystallization and data collection. Crystals were grown by hanging drop vapor diffusion at 24 °C over a reservoir solution containing 100 mM cyclohexylaminoethanesulfonic acid (CHES) pH 9.5, 950 mM

Table 1 Crystallographic analysis

Data collection statistics			
Crystal	SeMet MLN64 START domain		
Wavelength (Å)	$\lambda_1 = 0.9793$	$\lambda_2 = 0.9789$	$\lambda_3 = 0.9537$
d_{\min} (Å)	2.2	2.2	2.2
Completeness (%)	98.1	97.8	98.4
Observed reflections	136,560	139,509	141,249
Unique reflections	16,818	16,947	17,233
R_{sym}^1	0.065	0.074	0.072
Refinement statistics			
R-factor (%) ²	20.7		
R_{free} (%) ²	26.4		
Number of atoms	1,891		
R.m.s. deviations from ideality			
Bond lengths (Å)	0.017		
Bond angles (°)	1.85		

¹ $R_{\text{sym}} = \sum_h \sum_i |I_i(h) - \langle I(h) \rangle| / \sum_h \sum_i I_i(h)$; $I_i(h)$ is the i th measurement of reflection h and $\langle I(h) \rangle$ is the weighted mean of all measurements of $I(h)$.

²R-factor = $\sum_h ||F_{\text{obs}}(h)| - |F_{\text{calc}}(h)|| / \sum_h |F_{\text{obs}}(h)|$; R_{free} is calculated with 5% of the data.

articles

K/Na tartrate, and 200 mM LiSO₄. Crystals grew to a maximum size of 200 $\mu\text{m} \times 20 \mu\text{m} \times 20 \mu\text{m}$ over 2 days. Crystals were cryoprotected in reservoir solution supplemented with 20% (v/v) glycerol and flash frozen under N₂ gas at 95 K. The crystals are of space group P3₁21 with $a = b = 83.4 \text{ \AA}$, $c = 81.9 \text{ \AA}$. Three sets of inverse beam data (Table 1) were collected on an ADSC Q4 detector at beamline X4A at the National Synchrotron Light Source, Brookhaven National Laboratory. Data were processed with HKL³⁸.

Structure determination and refinement. Three selenium sites were located using automated Patterson search and cross validation in SOLVE³⁹. Phases calculated in SOLVE had a overall figure of merit of 0.67 (20–2.5 \AA), and 0.60 in the last resolution shell (2.58–2.50 \AA), and were improved with density modification in SOLOMON⁴⁰. The model was built into a 2.5 \AA electron density map with O⁴¹. Model refinement was carried out with CNS⁴² using torsional dynamics and the maximum likelihood target function and all measured data from the inflection point ($\lambda = 0.9793 \text{ \AA}$) set to 2.2 \AA resolution. The refinement was monitored using the free R-factor calculated with

5% of observed reflections. The refined crystal structure includes residues 230–443 of MLN64, 1 tartrate ion, and 165 water molecules. The structure has no residues in disallowed regions of the Ramachandran plot. The tunnel volume was estimated with SURFNET⁴³.

Coordinates. Coordinates have been deposited in the Protein Data Bank (accession code 1EM2).

Acknowledgments

We thank D. Stocco, C. Ponting and H. Bernstein for discussions; C. Ogata, R. Abramowitz, Z. Dauter, K. R. Rajashankar, J. Ho, and S. Misra for assistance with data collection at beamlines X4A and X9B, National Synchrotron Light Source, Brookhaven National Laboratories; A. Hickman for comments on the manuscript, and Z. Derewenda for providing expression vectors.

Received 14 February, 2000; accepted 21 March, 2000.

- Ponting, C.P. & Aravind, L. START: a lipid-binding domain in StAR, HD-ZIP and signaling proteins. *Trends Biochem. Sci.* **24**, 130–132 (1999).
- Schultz, J., Milpetz, F., Bork, P. & Ponting, C.P. SMART, a simple modular architecture research tool: identification of signaling domains. *Proc. Natl. Acad. Sci. USA* **95**, 5857–5864 (1998).
- Clark, B.J., Wells, J., King, S.R. & Stocco, D.M. The purification, cloning, and expression of a novel luteinizing hormone-induced mitochondrial protein in MA-10 mouse Leydig tumor cells. Characterization of the steroidogenic acute regulatory protein (StAR). *J. Biol. Chem.* **269**, 28314–28322 (1994).
- Sugawara, T. et al. Human steroidogenic acute regulatory protein: functional activity in COS-1 cells, tissue-specific expression, and mapping of the structural gene to 8p11.2 and a pseudogene to chromosome 13. *Proc. Natl. Acad. Sci. USA* **92**, 4778–4782 (1995).
- Lin, D. et al. Role of steroidogenic acute regulatory protein in adrenal and gonadal steroidogenesis. *Science* **267**, 1828–1831 (1995).
- Caron, K.M. et al. Targeted disruption of the mouse gene encoding steroidogenic acute regulatory protein provides insights into congenital lipid adrenal hyperplasia. *Proc. Natl. Acad. Sci. USA* **94**, 11540–11545 (1997).
- Waterman, M.R. A rising StAR: An essential role in cholesterol transport. *Science* **267**, 1780–1781 (1995).
- Stocco, D.M. & Clark, B.J. Regulation of the acute production of steroids in steroidogenic cells. *Endocrine Rev.* **17**, 221–244 (1996).
- Bose, H. S. et al. The pathophysiology and genetics of congenital lipid adrenal hyperplasia. *New Engl. J. Med.* **335**, 1870–1878 (1996).
- Miller, W.L. & Strauss, J.F. III. Molecular pathology and mechanism of action of the steroidogenic acute regulatory protein, StAR. *J. Steroid Biochem. Mol. Biol.* **69**, 131–141 (1999).
- Stocco, D.M. An update on the mechanism of action of the steroidogenic acute regulatory (StAR) protein. *Exp. Clin. Endocrinol. Diabetes* **107**, 229–235 (1999).
- Moog-Lutz, C. et al. MLN64 exhibits homology with the steroidogenic acute regulatory protein (StAR) and is over-expressed in human breast carcinomas. *Int. J. Cancer* **71**, 183–191 (1997).
- Watari, H. et al. MLN64 contains a domain with homology to the steroidogenic acute regulatory protein (StAR) that stimulates steroidogenesis. *Proc. Natl. Acad. Sci. USA* **94**, 8462–8467 (1997).
- Homma, Y. & Emori, Y. A dual functional signal mediator showing RhoGAP and phospholipase C- δ stimulating activities. *EMBO J.* **14**, 286–291 (1995).
- Raya, A., Revert, F., Navarro, S. & Saus, J. Characterization of a novel type of serine/threonine kinase that specifically phosphorylates the human Goodpasture antigen. *J. Biol. Chem.* **274**, 12642–12649 (1999).
- Masucci, J.D. et al. The homeobox gene *GLABRA 2* is required for position-dependent cell differentiation in the root epidermis of *Arabidopsis thaliana*. *Development* **122**, 1253–1260 (1996).
- Arakane, F. et al. Steroidogenic acute regulatory protein (StAR) retains activity in the absence of its mitochondrial import sequence: implications for the mechanism of StAR action. *Proc. Natl. Acad. Sci. USA* **93**, 13731–13736 (1996).
- Arakane, F. et al. The mechanism of action of steroidogenic acute regulatory protein (StAR). StAR acts on the outside of mitochondria to stimulate steroidogenesis. *J. Biol. Chem.* **273**, 16339–16445 (1998).
- Bose, H.S., Whittall, R.M., Baldwin, M.A. & Miller, W.L. The active form of the steroidogenic acute regulatory protein, StAR, appears to be a molten globule. *Proc. Natl. Acad. Sci. USA* **96**, 7250–7255 (1999).
- Gajhede, M. et al. X-ray and NMR structure of Bet v 1, the origin of birch pollen allergy. *Nature Struct. Biol.* **3**, 1040–1045 (1996).
- Gibrat, J.-F., Madej, T. & Bryant, S.H. Surprising similarities in structure comparison. *Curr. Opin. Struct. Biol.* **6**, 377–385 (1996).
- Flower, D.R. The lipocalin protein family: structure and function. *Biochem. J.* **318**, 1–14 (1996).
- Bose, H.S., Baldwin, M.A. & Miller, W.L. Incorrect folding of steroidogenic acute regulatory protein (StAR) in congenital adrenal lipid hyperplasia. *Biochemistry* **37**, 9768–9775 (1998).
- Kallen, C.B. et al. Steroidogenic acute regulatory protein (StAR) is a sterol transfer protein. *J. Biol. Chem.* **273**, 26285–26288 (1998).
- Westerman, J. et al. Identification of the lipid-binding site of phosphatidylcholine-transfer protein with phosphatidylcholine analogs containing photoactivatable carbene precursors. *Eur. J. Biochem.* **132**, 441–449 (1983).
- Garcia-Daguer, F.L. et al. NMR structure of the sterol carrier protein-2: implications for the biological role. *J. Mol. Biol.* **295**, 595–603 (2000).
- Reese-Wagoner, A., Thompson, J. & Banaszak, L. Structural properties of the adipocyte lipid binding protein. *Biochim. Biophys. Acta* **1411**, 106–116 (1999).
- Sha, B. & Luo, M. PI transfer protein: the specific recognition of phospholipids and its functions. *Biochim. Biophys. Acta* **1411**, 268–277 (1999).
- Derman, A.O., Puziss, J.W., Bassford, P.J. Jr. & Beckwith, J. A signal sequence is not required for protein export in *prfA* mutants of *Escherichia coli*. *EMBO J.* **12**, 879–888 (1993).
- Arakane, F. et al. Phosphorylation of steroidogenic acute regulatory protein (StAR) modulates its steroidogenic activity. *J. Biol. Chem.* **272**, 32656–32662 (1997).
- Hurley, J.H. & Misra, S. Signaling and subcellular targeting by membrane binding domains. *Annu. Rev. Biophys. Biomol. Struct.* **29**, in the press.
- Renaud, J.-P. et al. Crystal structure of the RAR- γ ligand-binding domain bound to all-trans retinoic acid. *Nature* **378**, 681–689 (1995).
- Williams, S.P. & Sigler, P.B. Atomic structure of progesterone complexed with its receptor. *Nature* **393**, 392–396 (1998).
- Shiau, A.K., et al. The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. *Cell* **95**, 927–937 (1998).
- Nolte, R.T. et al. Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptor- γ . *Nature* **395**, 137–143 (1998).
- Sheffield, P., Garrard, S. & Derewenda, Z. Overcoming expression and purification problems of RhoGDI using a family of parallel expression vectors. *Protein Expr. Purification* **15**, 34–39 (1999).
- Okamura, et al. A porcine homolog of the major secretory protein of human epididymis, HE1, specifically binds cholesterol. *Biochim. Biophys. Acta* **1438**, 377–387 (1999).
- Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
- Terwilliger, T.C. & Berendzen, J. Automated structure solution for MIR and MAD. *Acta Crystallogr. D* **55**, 849–861 (1999).
- Abrahams, J.P. & Leslie, A.G.W. *Acta Crystallogr. D* **52**, 30–24 (1996).
- Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).
- Brunker, A.T. et al. Crystallography & NMR system (CNS): a new software system for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
- Laskowski, R.A. SURFNET: a program for visualizing molecular surfaces, cavities, and intermolecular interactions. *J. Mol. Graphics* **13**, 323–330 (1995).
- Christopher, J.A. Spock. <http://quorum.tamu.edu/jon/spock/>